Journal Pre-proof

Part 1:- quality assurance mechanisms for digital forensic investigations: Introducing the Verification of Digital Evidence (VODE) framework

Graeme Horsman

PII: S2665-9107(19)30038-6

DOI: https://doi.org/10.1016/j.fsir.2019.100038

Reference: FSIR 100038

To appear in: *Forensic Science International: Reports*

Received Date: 15 August 2019

Revised Date: 6 September 2019

Accepted Date: 9 September 2019

Please cite this article as: { doi: https://doi.org/

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.

Journal Pre-proof

Part 1:- Quality assurance mechanisms for digital forensic investigations: introducing the Verification of Digital Evidence (VODE) framework

Graeme Horsman

Teesside University Middlesbrough Tees Valley TS1 3BX

Email: g.horsman@tees.ac.uk

Abstract
Quality assurance measures in the field of digital forensics play a vital role for upholding and developing investigatory standards. Coupled with the fast pace of technology, practitioners in this discipline are often faced with the challenge of interpreting previously unseen or undocmented forms of potentially evidential digital data, content which may be crucial to a current case under investigation. Mechanisms to support this interpretative process offer support for the practitioner, helping to guide them through this task and the steps involved in ensuring any reported information is accurate. This work presents the Verification of Digital Evidence (VODE) framework, designed to support digital forensic practitioners when testing and verifying their interpretation of digital data. The stages of VODE are discussed and its application placed in context.

Keywords:- Quality Assurance; Digital Forensics; Digital Evidence; Forensic Science

1 Introduction

As scrutiny of the forensic sciences continues, the sub-discipline of digital forensics (DF) has not escaped its share of criticism (Pollitt et al., 2018; Science and Technology Select Committee, 2019; The Guardian, 2019; Daeid et al., 2019; Page at al., 2019c). As one of the newer sub-disciplines of forensic science, albeit now well established as a vital tool of criminal justice systems world-wide, DF still requires the development of robust and viable methods to support practitioners in their investigation of digital exhibits and the production of reliable evidence. The DF field has received much well-documented coverage of its issues and therefore this article will not repeat already well covered ground, instead, offering a synopsis of these. A lack of standardisation and consistently applied examination techniques (Mohay, 2005; Grobler, 2010; Valjarevic and Venter, 2012; Kebande and Ray, 2016; Lillis, et al., 2016), practitioner accreditation and education (Cooper, et al., 2010; Sommer, 2011; Sommer, 2018), and cost and resourcing cuts have all lead to varying levels of quality of DF casework and concerns with the work produced by this field.

There is no doubt that as with other branches of forensic sciences, DF maintains a number of problem areas which are in need of being addressed, as noted in the recent Science and Technology Select Committee's (2019) examination of forensic science in the United Kingdom. Addressing many of these issues will take time, requiring sustained investment and development of infrastructure, knowledge and services (Gov.uk, 2017), a debate beyond the scope of this paper. Whilst effective financial investment in DF has the potential to improve

Journal Pre-proof

standards, there are areas of quality assurance in DF which arguably require procedural changes in order to improve standards. In DF, one of these areas involves the processes a practitioner undertakes to determine the reliability of their interpretation of evidence which they present as part of their case work, in other words - '*how a practitioner tests and verifies their interpretation of any potentially relevant digital data*'. This will remain the focus of this work.

Frequently research focuses on the development of frameworks and schemas to tackle specific technology types (see work by Carrier and Spafford, 2004; Martini and Choo, 2012 with regards to cloud computing; Petroni Jr et al., 2006 for volatile memory; Chung et al., 2017 for home assistants). Whilst this has value for the field in determining approaches to tackle this content when it is encountered by a practitioner (Horsman, 2019b), often the need to develop approaches designed to help a practitioner reliably interpret, and subsequently verify this interpretation of digital data is overlooked, despite being a fundamental task in all investigations. Few inroads within this area have been made, where Pollitt et al's (2018) contribution via The Organization of Scientific Area Committees for Forensic Science (OSAC) remains one of the only attempts.

In all cases requiring a DF examination, a DF practitioner will sift through a digital device for potentially relevant information which can support the criminal justice system to effectively reach a decision with regards to whether an offence has been committed. A practitioner's examination of data will typically attempt to establish one of two things, either the occurrence or non-corruence of a specific set of actions, or to identify or attribute behaviours to a specific or subset of users (see similar concepts in DNA analysis (Kokshoorn et al., 2017). This process requires the practitioner to determine the relevance of digital information and present content that supports the accurate reconstruction of events which were alleged to have taken place. Any misinterpretation of information at this stage creates uncertainty, inaccuracies and ultimately prevents any reported information from the practitioner from being relied upon. Of concern, lies the issue that if such misinterpretation occurs, this may never actually be detected or noticed.

Whilst tech-specific frameworks may be applicable to only a subset of a practitioner's overall caseload, practitioners should interpret digital data in every case they undertake, emphasising the need for support in this aspect of their role. The changing DF landscape now arguably favours options for quickly processing information with automated approaches (Horsman, 2019d), where it is easy to simply report the output of scripted tools as verbatim, without applying any additional level of validation. While this article does not suggest existing tool quality is poor, in fact quite the opposite, where practitioners have access to some of the best and skillfully created tools to date, it does however stress that the interpretation of a tool's output must be applied by the practitioner. Where there is no interpretation, the practitioner is merely reporting a tool's results, in the absence of case and surrounding system metadata context. The addition of a reliable interpretation of data allows for the circumstances of a case to be reported more robustly (Horsman, 2019d) (see Figure 1).

The interpretation phase of an investigation poses a challenge to the practitioner, and there is little formalised guidance and support for this task in DF. There is a notable lack of DF-specific

Journal Pre-proof

published material which is designed to support a DF practitioner test and validate hypotheses, describing how to do this and what is involved in this process. This work aims to provide additional interpretative assistance to the practitioner and offers the Verification of Digital Evidence (VODE) framework to support their interpretation of newly encountered, undocmented digital artefacts and data. This work forms the first of two pieces concerning measures for quality assurance in DF investigations and will bring together four frameworks which are designed to support the DF investigatory process (see Figure 2). The proposed VODE framework, the Framework for Reliable Experimental Design (FRED), the Digital Evidence Reporting and Decision Support (DERDS) framework and the Capsule of Digital Evidence (CODE) (proposed in Part 2 of this submission) combine to support the practitioner with their application mapped in Figure 2.

2 The Verification of Digital Evidence (VODE) framework

From the outset it is necessary to define the role of the VODE framework. The VODE framework is there to support a DF practitioner when they have found information in a suspect case and they need to interpret its meaning and subsequent impact on the investigation. The practitioner will have the 'baseline' facts of a case which include reference to a suspected offence type, and the set of actions which have triggered the need for an investigation. During an examination of target media, potentially relevant digital content may be flagged through initial data searching and processing procedures, and following identification, the practitioner must attempt to determine its meaning, connection to the alleged offences and those actions which could be responsible for its presence on a system. In essence, the practitioner is attempting to accurately determine what digital events have or have not occurred through the interpretation of digital traces (defined below in Section 2.1), where support from VODE is offered. The VODE framework is not designed to help the practitioner to find potentially relevant data on a system. Further, VODE assumes those practitioners engaging with it are suitably competent to carry out their role.

Before a complete discussion of the VODE framework is offered, it is first presented in Figure 3 for review. A subsequent discussion of the elements VODE comprises of is provided in the following sections.

To provide some context, the VODE framework is designed to support those practitioners engaging with the verification of their interpretation of '*new knowledge*' as part of a DF investigation. As a result, it is necessary to discuss what '*new knowledge*' is within this context.

2.1 What is a '*new knowledge*' scenario?

Within the context of DF this work coins a '*new knowledge scenario*' as:-

> *'A new knowledge scenario denotes a situation where a practitioner is encountering a digital artefact/data or case scenario for the first time,* ***AND*** *it is also previously undocumented by anyone* ***OR*** *an interpretation exists but it is unreliable. In such a case, a practitioner must interpret this information using robust scientific methods to provide accurate meaning and context.'*

Journal Pre-proof

New knowledge scenarios in DF involve the interrogation of digital traces, defined by Pollitt et al., (2018) as 'vestiges left from a past event or activity, criminal or not...a trace is any modification, subsequently observable, resulting from an event.'. Traces may either be fragments of digital data recovered during carving processes of actual digital artefacts which exist on a system, defined by Horsman (2019b) as 'a digital object containing data which may describe the past, present or future use or function of a piece of software, application or device for which it is attributable to'. Regardless of the trace-type, a practitioner must generate a reliable interpretation of it themselves, based on their own investigation of it.

The product of a '*new knowledge scenario*' is '*new knowledge*', generated by the practitioner following their own testing and interpretation of a digital trace. This occurrence may be frequently encountered in the DF field due to the rapid pace of technological change, and therefore it is likely that all practitioners will at some stage in their career encounter a digital trace which they must interpret without any existing guidance. The importance of approaching '*new knowledge*' occurrences effectively cannot be understated, as not only may a direct case featuring such data be impacted by misinterpretation, further distribution of such knowledge by a practitioner may lead to an additional negative impact on a wider scale. When a '*new knowledge*' scenario occurs, requiring reliable interpretation, those involved with this process must understand the steps required to approach this situation and to reliably interpret the digital traces which they believe is relevant to their case.

One of the first challenges for a practitioner is to determine whether they are actually experiencing a '*new knowledge*' scenario, and support for this process is offered by the Digital Evidence Reporting and Decision Support (DERDS) framework (Horsman, 2019). The DERDS framework provides a defined route for a practitioner to follow for assessing whether a given digital trace found within their current examination has previously been encountered and reliability interpreted - which they may be able to rely upon themselves for their current examination. This includes three pathways, first where the digital content in question has already been interpreted as part of a past case, second, where interpretation has been provided in published and peer reviewed literature, and finally the third, where a new knowledge situation is present. Crucially, for both the first and second pathways, the reliability of any source forms a key criterion, and therefore the practitioner must assess whether any past interpretations of data can be relied upon, taking into account factors such as the source of any interpretation, whether it has been subject to peer review and whether there is transparency and thoroughness in any testing which was carried out, which has ultimately lead to the produced results. If previous interpretations are available and reliable, then the DERDS framework permits the use of this content, creating an '*existing knowledge*' scenario. In the third instanced, the VODE framework is designed to support the practitioner.

Where a '*new knowledge*' scenario is present, the practitioner is required to use their existing skills and knowledge to determine a reliable interpretation of target digital data. This ultimately requires a series of procedural steps which includes testing to correctly establish a set of given facts.

Journal Pre-proof

2.2 Initial hypothesis / assumption generation

Assuming that the DERDS framework has been followed and the practitioner is in a position where the data from a suspect system requires their interpretation and the verification of '*new knowledge*' is required, engagement with the VODE framework can continue and the following initial questions should be addressed.

1. *How have I found the artefact/data?*: This is a key question which prevents reliance on any potentially erroneous underlying procedures which may have been used during the initial searching and recovery of an exhibit's content. Here the practitioner should be able to determine that any 'potentially relevant' digital traces have been found via processes which have been run correctly, completed correctly and have targeted the correct case exhibits. Further, the practitioner should understand the process they have ran, preventing reliance on push-button forensic procedures, where understanding of the process is lacking (Baggili and Breitinger, 2015). It is important to note at this point, that tool testing and validation is considered an external factor beyond the scope of VODE and it is assumed that all practitioners have tested their tools and they are functioning correctly.

2. *Why have I found the artefact/data?*: The practitioner must understand why they have found the 'potentially relevant' content in question on a suspect device. At this stage, understanding only extends as far as acknowledging why an initial discovery has been made, whether by a keyword search, carved content or as the result of a bespoke recovery script supplied by a specific tool vendor. A practitioner should not be in a position where they have no understanding of why they have come across any of the digital traces in question. Practitioners should have a record and understanding of all implemented forensic processes and can attribute any outputs to their respective process.

3. *What do I think it means?*: In order for any data to be deemed initially potentially relevant, a practitioner must have imparted some form of investigative instinct upon this information. The formalisation of initial suspicions regarding the data helps to establish the potential worth of it in regards to the case in question, following its reliable interpretation. It also helps to formulate and develop testing processes to be implemented later in the VODE framework. If a practitioner believes that data is linked to a specific process and this action would be evidentially relevant, then testing should be designed around exploring this in order to accurately confirm or refute this theory.

4. *What is the potential case impact?*: Establishing case impact helps to determine worth. In the current climate, practitioners are more stretched than ever with regards to case processing timeframes, limited resources and organisational backlogs. Arguably there is no room for the exploration of non-relevant digital content and where the digital trace's potential case impact is low or non-crucial, a decision to not implement further testing may be necessary. For example, where prosecution thresholds are in place and have

Journal Pre-proof

already been met, further work may not be necessary (regardless of ethical and moral debates raised by such approaches).

5. *What has the suspect potentially done to create this artefact/data?*: The final question for a practitioner to consider is an amalgamation of the previous four. The practitioner should consider what a suspect has potentially done on their system in order to have generated this potentially relevant content, and why this is important to the case in question. These actions should be considered during the generation of test cases to determine whether such behaviour could have occured.

The creation of initial hypotheses and assumptions helps to provide focus and meaning to the subsequent testing of these digital traces which must follow as part of the interpretation process. In addition, consideration of the digital facts which concern the digital traces is also required.

2.2 Establish digital facts surrounding the artefact / data's presence

Whilst generating initial hypotheses in the aforementioned stage focuses on offence-based and procedural circumstances, the practitioner must also consider any digital facts attributable to the potentially relevant digital traces. The following should be contemplated.

1. *What processes do I think caused the creation of the artefact/data?*: Establishing the presence of such data is only the first step of the investigatory process. Given its location on a system and the processes which ultimately flagged this content to the practitioner, initial suppositions regarding what is believed to have caused the presence of this data on a suspect device should be generated. In doing so, this forces the practitioner to consider the surrounding circumstances of the case, content which is also present on the device, and the actions believed to have been carried out on a system. It also supports the practitioner in determining whether the data is in fact potential evidentially relevant. Data may have been directly created by a piece of software under suspicion (for example, a chat client which has created a chat log in a grooming offence), or via passive operating system processes which have recorded user-system interaction.

2. *Where is the artefact/data and what does this mean?*: Determining the location of data on a system or device and what this location means, provides the practitioner with initial intelligence. Whilst they may not be in a position to determine the relevance of the data itself, file system location information can provide interpretative support. Many of the main operating system structures have now been well documented, and therefore the location of data may support what a practitioner thinks in terms of its relevance to their case.

3. *What do I think the artefact/data is?*: File signature information or internal structural content may provide some interpretative support. For example, common structures such as SQLite are often used in certain platforms for application logging. If a structure can be identified, then any relevant tools can be identified for parsing and displaying data in

Journal Pre-proof

order to review relevant content. Unknown structures require a greater amount of testing as 'viewer programs' may not exist and data must be manually extracted and displayed.

4. *Why do I think the artefact/data is there?*: Here the practitioner should consider why the data has occurred on the device in question and what digital actions are believed to have been responsible for generating this information.

5. *Establish any surrounding applicable metadata for artefact/data*: The practitioner should identify associating metadata which may be linked to artefact/data which describes its function on the suspect system. This could be file system metadata following a set of actions or operating system logs and artefacts which capture data passively when the system is running. The reliability of this information must also be considered and whether such data relates solely to the digital traces under consideration or whether multiple actions may cause its manipulation.

Once all hypotheses have been generated and digital facts established, the testing processes can begin, commencing with deciding upon the type of reconstructed test environment which will be used.

2.3 Reconstruction Phase

As defined by Pollitt et al., (2018, p.7) reconstruction a suspect digital environment involves the organisation of 'observed traces to disclose the most likely operational conditions or capabilities (functional analysis), patterns in time (temporal analysis), and linkages between entities – people, places, objects – (relational analysis).'. A reconstructed digital environment provides the foundation for any following testing to occur.

Digital traces on device occur as a result of either a singular or set of actions and where these actions appear evidentially relevant, and with most forensic science evidence types, the practitioner must attribute these to a specific action or actor (Kokshoorn et al., 2017). Therefore where potentially evidential content has been found by a practitioner, its presence is unlikely due to chance, and more likely the result of a direct action or actions on that device; an instance of digital cause and effect. In essence the practitioner must explore which actions can be responsible for the presence of the data of interest on a device, and this is often achieved through testing carried out in a reconstructed environment, mimicking that environment experienced by a suspect.

The reconstruction of digital events for the purpose of testing is a difficult but necessary task where a practitioner must recreate the believed actions of the suspect as closely as possible to establish what occured on the suspect system. Given the standard required of forensic science evidence, reconstruction must be effective and reliable. Reconstruction is defined as '*an attempt to get a complete description of an event using the information available, or an attempt to repeat what happened during the event*' (Cambridge Dictionary, 2019). A reconstruction of digital events should be complete, matching the environment used in any alleged offence in order for

Journal Pre-proof

any resulting interpretation to be reliable. In such cases, the practitioner should consider variables such as the following:

1. The exact version of the application/software responsible for the data.
2. The exact version of the operating system.
3. Comparable hardware (if this is a factor for consideration).
4. The exact suspected test actions on the application/software responsible for the data. For example, comparable transfer of communications or file types.

The reconstruction phase will almost always pose a challenge and it must be stressed that it does form a vital part of the DF process. Any reconstruction which falls short of complete (for example, mismatched operating system or software versions are used), provides a level of functionality doubt, and ultimately undermines the reliability of any subsequent testing and results (see Figure 4). Unlike more physical actions requiring traditional body fluid or trace evidence forensic procedures, practitioners in DF are in a unique position in that they are in a better position to carry out a complete reconstruction of suspect scenario, addressing the four points noted above. The aim of reconstruction is to simulate the actions of the suspect, using the same circumstances available to the suspect at the time of an alleged offence. Any discrepancies in this setup raise questions regarding the applicability of any subsequent interpretation of data following testing as the circumstances were dissimilar. Despite this, questions of necessity will be raised, and these should be addressed.

Is a complete reconstruction necessary?: Complete reconstruction is a heavy burden for the practitioner to carry, and questions of necessity may be raised. In a period where DF organisations are facing pressures to process content efficiently and effectively, adequate reconstruction and testing is an additional labour intensive process to bear, requiring time to complete properly. The issue here is the impact of testing which is not completely applicable to data found on a suspect system. Whilst complete reconstruction may be a challenge, it should be seen as the gold standard, and achieving this removes any questions of doubt regarding subsequent results. Given DF is a tool for criminal justice systems to deliver robust, reliable justice based on reliable evidence, complete reconstruction and testing is arguably a must. However, this also raises the question - '*what happens when complete reconstruction is not possible?*'.

Where discrepancies in the reconstruction phase occur due to an impossibility for complete reconstruction or the process is impractical, an impact on the reliability of subsequent testing and interpretation is inevitable. This creates a 'partial reconstruction' situation, where the following points are raised.

1. *If complete reconstruction cannot be practically achieved, what is the next highest level of reconstruction that can be achieved?*: It would be naive to imply that complete reconstruction will be an option in all circumstances (despite the fact it may be achievable in a high proportion of cases). Factors such as a lack of resources, lack of knowledge and lack of time, despite not being acceptable, may still be a barrier. Where it

Journal Pre-proof

is not achievable, a reconstructed environment should attempt to be as close to the initial scenario as possible. In some cases, '*closeness*' may be a case of simply testing an older iteration of an application or software. Yet such an approach should be treated with caution and arbitrarily utilising metrics such as a software version to define the closeness of a reconstructed environment can be misleading given potential wholesale changes which can occur with software upgrades and releases. As a result, defining the closeness of a reconstructed environment becomes multifaceted. Factors to consider should include content such as a software version number, but should also consider whether the functionality of an app (or sub-function of an app) is the same, allowing a user to undertake the same actions in the same way and that underlying logging appears consistent in terms of location, structure and internal metadata types.

2. *Can I explain the differences and potential impact of the reconstructed environment and potential impact on my interpretation of data on the suspect system?*: If discrepancies in the test environment are apparent, a practitioner must be able to explain their presence and impact, and ultimately, where their testing allows them to reliably interpret the digital trace found on a suspect system.

3. *Can I rely on my interpretation of data given discrepancies in the reconstructed test environment?*: Crucially, here the practitioner must accept that the results of their testing cannot be 100% applied to the case under investigation as the circumstances of the original suspect data and the testing are different, regardless of the (unlikely quantifiable) amount. The challenge then lies with determining whether results can be relied upon to then apply to the practitioners current case and subsequent interpretation of data.

When attempting to reconstruct a suspect environment for testing there are two options, either 'clean' of 'dirty'.

Clean Reconstruction: A clean reconstructed environment involves the use of a fresh installed system (comparable operating system and version), either installed on a clean disk drive to boot with physical hardware, or a virtualized version. In either case, the base operating system is untouched, where all resulting actions following testing belong to the suspect software which is installed upon it, the operating system, and the set of test actions which a practitioner has carried out. This helps to prevent test contamination and limits the background noise that can occur on a system and make the interpretation of results more difficult. When the clean operating system is configured, the exact version of suspect software should be downloaded or installer files could be extracted from a suspect system image and installed.

Dirty Reconstruction: A dirty reconstruction involves the use of the suspect system itself and involves the reverse image of a suspect device onto a secondary drive, or a virtualisation of the suspect system using software such as VFC (MD5, 2019). This form of reconstruction places the practitioner in the position that the system was in when last used by a suspect, where all previous data created by a suspect's actions are available to the practitioner for scrutiny.

Journal Pre-proof

2.4 Testing

Testing is undertaken to assess and establish the correct functionality, behaviour of a particular artefact or set of data on a system and to establish what actions have caused its presence. In this context, testing must involve the replication of those actions believed to have been carried out by a suspect on their system, on a comparable test environment. The results of these tests must be analysed for meaning, and this interpretation applied to data found on the suspect system for the purpose of reliably confirming or denying what a suspect is alleged to have done. To support the testing process, the Framework for Reliable Experimental Design (FRED) (Horsman, 2018) should be utilised. Here the necessary stages for carrying out effective testing of digital content for the purposes of forensic interpretation are defined.

A FRED reminder: The FRED framework maintains the six core stages (planning, implementation, evaluation, repetition, analysis and confirmation) required for developing and undertaking robust testing of digital artefacts/data. Practitioners should follow FRED when planning and implementing their testing program. In addition, support for interpretation of results is offered (Horsman, 2018).

Often testing involves trial and error, guessing what a suspect may have done based on initial case-sight, trialing these actions on a test reconstruction, and determining whether the results of these actions are comparable to those on the suspects device. Even where test actions match those present on a suspect system, this alone does not necessarily provide confirmation of a suspect's behaviour and the consideration of alternative actions must be factored into testing (see further discussion in Section 2.4.2). Where possible, testing must aim to establish that only a set of certain actions can result in the data traces present on the suspect device. For example, a practitioner must make sure that action ‘A’ and only action ‘A’ creates result ‘B’ on a system. Of course, this will not be achievable in all cases, where digital traces may be present due to one of a number of actions, which may not be distinguishable on the suspect device. Therefore, the practitioners must make sure that they don't mistake what is coined in this work as a ‘possible but unconfirmed’ result for one which they believe is ‘confirmed’ (see Figure 5). Often trial and error testing is undertaken, where a practitioner must explore each individual function of an application/software suspected.

In most instances, practitioners will not engage with source code analysis of any software/application believed to have caused the presence of potentially evidential digital traces on a suspect machine. This will arguably be either due to a lack of availability, or a lack of time, resources and knowledge to carry out this form of testing. Instead, ‘black box’ styled testing will take place. Black box testing involves assessing an application’s behaviour by utilising a set of known inputs to generate and analyse outputs without ever knowing the internal structure and function of the software itself (Nidhra and Dondeti, 2012). As a result, utilised inputs must be comprehensive enough to fully exercise the tested software’s functions. This is a difficult task as without understanding internal code structures, a practitioner must use a combination of intuition, guesswork and the visible functions that the software offers the user in order to generate a comprehensive set of inputs (see Figure 6). This may also involve a practitioner

Journal Pre-proof

using the suspect software in order to understand what it is capable of, then isolating and testing specific functionalities of it.

The risk of black box testing in this context is twofold:

1. *Weaknesses in inputs*: The black box testing process relies on the generation of test inputs which will exhaust the functionality of the software (or a target function of the software), generating all potential output sequences. In doing so, the practitioner aims to be in a position to understand how the software behaves in all input situations. Ultimately, this will also reveal the behaviours of the software which result in data which is comparable to that found on the suspect system. At this point a practitioner can determine what a suspect must have done on their system to have caused the presence of any potentially evidential content. Getting to this point requires rigorous testing to exhaust all possible black box outcomes. Failure to exhaustively test leaves potential unknown functionality which may have been responsibility for suspect behaviour but remain unknown to the practitioner. Conversely, as shown in Figure 5, where multiple actions may be responsible for the same outcomes, failure to establish this may cause unreliable interpretations of suspect data.

2. *Misinterpretation of output*: Misinterpretation of outputs when black box testing remains a risk. Where analysis of source code can produce an objective description of code functionality, in black box testing, the practitioner must subjectivly interpret what the outputted data means in conjunction with the utilised input actions, whilst also considering the possibility that further inputs may also result in the same output.

Due to the nature of black box testing approaches, testing must be a logical, iterative process, designed to assess one functionality of a piece of software at a time to reduce the chance of misinterpretation. This should be taken into account when determining the input criteria used during each test. For each test carried out, all criteria used and resulting software outputs must be documented, repeated and archived. In addition, any resulting interpretations of data must also be recorded and archived.

2.4.1 Attaining a sufficient level of testing

Testing in this capacity can be a resource-intensive task raising the question '*when have I undertaken sufficient testing in order to be able to reliably interpret the data highlighted on the suspect system?*'. In order to answer this question, the initial purpose of testing in the first instance should be consulted. In essence, testing is taking place due to the need to establish the reliability of the interpretation of '*new knowledge*', discovered as part of an active case under investigation. In order to do this, testing must put the practitioner in a position where they can fully describe the artefact/data and it's structure and functionality. Only when this has been achieved, has sufficient testing been carried out, as it is only at this point would a practitioner have all the necessary test data to be in a position to interpret the data in question reliably. Testing which falls short of allowing an artefact/data to be fully described leaves potential unknown functionality undiscovered which may impact an overall interpretation of data. The

Journal Pre-proof

practitioner must also consider factors of repeatability, and test actions and results must be constantly repeatable to remove the chance of misinterpretation occurring from a 'one time event'.

2.4.2 Considering alternatives scenarios

When testing has described an artefact/data's functionality, alternative scenarios must also be considered (Horsman, 2019) as noted above in Section 2.4. Here, consideration of competing hypotheses enables a thorough evaluation of any testing processes to take place, making sure that all potential scenarios have been tested and outcomes evaluated against original practitioner hypotheses (Casey, 2018). In some digital scenarios, multiple actions may result in the same output and each must be tested and verified within the confines of the suspect system data. It may be easy to fall into the trap of relying on the first set of tests which provide an outcome matching an original hypothesis. By considering actions which may also be responsible for comparable outcomes, two issues may be raised:

1. *Requirements for further testing or further developed tests*: Further testing or developed tests may be required to determine reliably what has occured on a suspect device following consideration of competing hypotheses. When considering competing hypotheses, where additional explanations may exist for explaining the presence of a digital trace on the suspect system, these must be explored. A practitioner must establish the validity of competing hypotheses via either further test implementations or through the development of additional tests which directly target any alternative scenarios.

2. *Inconclusivity*: Where following testing, an alternative hypothesis proves potentially valid, this may reveal that in the original case, a lack of available data on the suspect machine exists, prohibiting the practitioner from conclusively determining a suspect event. It is important that a scenario describing a set of data on a suspect machine is not presented as fact where uncertainty exists. Further, it must be recognised that digital data on a system does not always permit the factual depiction of events in all circumstances, and in some cases a lack of data may exist, creating uncertainty. In such cases, this must be acknowledged by the investigating practitioner.

2.5 Interpretation of findings

As Pollitt et al., (2018, p3) state, 'the value of forensic science as a whole is that it uses scientific reasoning and processes to address questions specific to an event or a case – for legal contexts, to provide decision-makers with trustworthy understanding of the traces in order to help them make decisions'. This process includes accurately interpreting the results of testing, a crucial and complex task. Any misinterpretation is both dangerous for the current case under investigation (where erroneous verdicts may be reached) and where a misinterpretation has been shared, any future investigation which has relied upon this to interpret a comparable digital trace. Further, as Pollitt et al., (2018, p3) note, 'there is the potential for cognitive bias and other non-technical sources of error to skew forensic results.', an area expanded on by Sunde and Dror (2019).

Journal Pre-proof

The interpretation of results should be a three step process. First, test results must be interpreted by the practitioner themselves. Second, results must be interpreted by a suitable third party practitioner (external to the original case) and compared to the primary practitioners interpretation. Where discrepancies are apparent, further testing should be undertaken. The utilisation of a third party practitioner to validate an interpretation who are independant to the investigation and case facts help to prevent interpretative bias being imparted (Pollitt et al., 2018). If testing has resulted in the peer review and confirmation of results by both practitioners, arguably best practice dictates that the third stage is the production of a standard operating procedure (SOP) for the interpretation of the artefact/date in question. The produced SOP not only ensures helps to confirm the practitioner understands the work they have done by being able to document it, but also provides a blueprint for those encountering this data-type within a prospective investigation, permitting future scrutiny, review and development of any documented processes. In addition, the SOP helps to promote consistency in interpretation in future cases, preventing misinterpretation and increasing investigation efficiency by preventing re-testing of the same scenario.

2.6 Apply interpretative methodology to original artefact / data

The final stage of the VODE framework is to apply the results of testing and the interpretation of findings to the data present on a suspect machine. Here, any parsing methodologies which have been developed or structural blueprints for the data/artefact should be applied to an extracted copy of the suspect data and results examined. If testing has been effective, any applied methodology should provide a reliable description of what events have occurred on the suspect device which were ultimately responsible for the presence of the potential evidential data. These results should also be validated by an appropriate third party who can verify that the methods developed during testing have been correctly applied to the suspect data and that any concluding interpretation is valid. This element of peer-review supports the robustness of investigatory process (Page et a., 2019c). Following completion of this stage, the practitioner has arguably undertaken the necessary steps involved in reliably interpreting digital content on a suspect system in a scenario requiring the interpretation of '*new knowledge'*. These processes should be captured and archived for the purposes of reuse and sharing, and therefore the concept of a Capsule of Digital Evidence is introduced.

3 Capsule of Digital Evidence (CODE)

This work partners with '*Part 2:- Quality assurance mechanisms for digital forensic investigations: knowledge sharing and the Capsule of Digital Evidence (CODE)'*. The process of reliably interpreting digital evidence is paramount to DF and the criminal justice system who seeks to rely upon it. The VODE framework is designed to help attain evidential reliability within the case work completed by DF practitioners as a quality assurance mechanism. CODE is a schema designed to capture the processes undertaken when a practitioner engages with VODE, so that this information can be archived and distributed. CODE submissions can be housed centrally for use in future cases within an organisation, as well as distributed between others within the field of DF. Doing so increases knowledge sharing and encourages its peer

Journal Pre-proof

review. CODE provides a field-wide collaborative mechanism, designed to support the quality of work undertaken.

4 A worked example of VODE

Now that the VODE structure and use has been presented, a worked example of its application can be provided. For this purpose, the example concerns the identification of a URL in the unallocated space of a device which leads to a site maintaining illegal content. This example will hypothetically assume that simply visiting this site is an offence, and therefore proof that this has occured on the suspect machine is an initial purpose of the investigation.

Scenario digital trace summary: The data '`www.<examplefictionalillegalsite>.com`' has been identified in the unallocated portion of the suspect's laptop device hard drive. This directs the user to a live website containing illegal material.

Consideration of DERDS: As VODE only concerns situations where a digital trace/scenario has never previously been encountered or reliably documented, the practitioner must engage with the DERDS framework to establish whether any existing published and peer reviewed work exists or past case precedents where a similar scenario has been examined. If this material exists, DERDS will guide the practitioner to decide whether this information can be used to help them establish the surrounding facts regarding the presence of '`www.<examplefictionalillegalsite>.com`' on the suspect system. If not, the practitioner must proceed with the stages of VODE.

Initial hypothesis: The presence of '`www.<examplefictionalillegalsite>.com`' in the unallocated regions of a device can be due to a number of activities and this must be considered. In this scenario, the practitioner has found the data as a result of a keyword search for a '*term of interest*' worklist. The data appears to be in the structure of a URL. Based on this data, the assumption is that the suspect has visited an illegal website using this device and subsequently deleted their browsing history. Therefore it is assumed this data originates from an Internet browser, but there are no live browsing history records indicating the browsing of illegal websites. The impact of establishing this information is the identification of a potential offence having been committed.

Establishing the digital facts: In this case there is a singular URL structure where it is difficult to determine origin. Structurally, the data appears to be a URL (where the chance of randomised data forming a valid URL to an active site is unlikely). This data was highlighted following a keyword search for known terms of interest. As only the apparent structure is initially identified, the practitioner should examine content surrounding the data in unallocated in order to establish the existence of potentially related metadata. For example, the URL may either be part of a carvable log file which exists in the deleted space of the drive, or that it simply maintains additional associated metadata describing the URL potential visit. This data may also be in a format which allows its origin to be identified - for example, metadata may be in a format which is attributable to a specific software application, or in this case a web browser. A practitioner should also identify any possible applications present on the device which are potentially

Journal Pre-proof

capable of generating this information. For the purpose of this example, Mozilla Firefox is the only browser installed and the only apparent source of this data (no other browsers or Internet related software is present).

Reconstruction: In this case, the practitioner must test the functionality of Mozilla Firefox in the same circumstances as the suspect device in order to determine whether it could be responsible for the apparent URL. The same version of the suspect device operating system and exact version of the Firefox browser must be tested. These are installed on a clean drive which is booted in a physical device.

Testing: Once the reconstructed environment is ready, the practitioner must design and implement tests designed to recreate the presence of a URL structure in the unallocated region of a device - comparable to that witnessed on the suspect device. The practitioner must identify functions of the browser and potential suspect actions which they believe could have created this scenario. Using test unique websites as visit criteria (so they can be identified as test actions), test cases must be carried out iteratively and the results analysed and compared. Engagement with the FRED framework will support the testing process. As the apparent URL is in the unallocated regions, a practitioner should examine the Firefox browser's functions which concern browser history deletion as well as any other methods a user could have utilised to delete their records. Testing must continue until a practitioner is confident that they have identified a set of actions which are responsible for the data present on the suspect device, or alternatively, that a set of actions cannot be reliably determined.

As part of this process, alternative hypotheses must be considered, such as actions like the apparent URL being present on the system due to it being passively acquired as part of a legitimate website's web page being cached. All potential avenues should be considered.

Interpretation: When testing is complete, the findings must be interpreted by the practitioner to determine their understanding of the actions carried out and subsequent results generated via testing. This must also be undertaken via a suitable third party practitioner who is independent to the investigation. Both cases should consider the original generated hypotheses and digital facts established earlier. If both the partitioner and reviewer are in agreement, the practitioner should formalise this process into an SOP.

Application of interpretation: The agreed practitioner test results must be applied to the data found on the suspect device (the unallocated '`www.<examplefictionalillegalsite>.com`') in order to attempt to accurately explain its presence. This should also be confirmed via a suitable third party practitioner who is independent to the investigation. If in agreement, a practitioner should be in a position to reliably explain those actions which are responsible for the apparent URL. It should be noted that this position may also include determining that there is insufficient evidence to conclusively identify those actions which are responsible for the URL. In either case, the practitioner should have carried out these processes in a way which minimises the potential for misinterpretation.

Journal Pre-proof

All steps of this process should be captured and submitted to CODE.

5 Conclusions

Given recent concerns over quality assurance processes for DF evidence, this work provides a twofold contribution:

1. The VODE framework is offered to support practitioners when interpreting digital evidence which has not previously been encountered (a '*new knowledge*' scenario). VODE provides a formalised method, noting all the stages which should be undertaken by a practitioner when attempting to explain the meaning of a digital trace found on a suspect system in relation to a set of suspected actions believed to have occured in their case. This work partners submission 'Part 2:- Quality assurance mechanisms for digital forensic investigations: knowledge sharing and the Capsule of Digital Evidence (CODE)'.
2. This article has provided a road map for the use of the author's four quality assurance frameworks (DERDS, FRED, VODE and CODE).

As the liklyhood of practitioners encounter previously unknown and undocumented digital structures, there is a need for mechanisims to support practitioners exmaine and interpret this data reliably. The VODE framework provides procedural guidance for the practitioner and encourages engagement with all the elements required when interpreting digital data through engagement with testing. VODE is the precursor to the CODE schema offered in Part 2, where VODE data can be captured and shared.

Conflicts of interests

There are no conflicts. I am the section editor of the digital forensic theme.

Journal Pre-proof

Journal Pre-proof

References

Baggili, I. and Breitinger, F., 2015, March. Data sources for advancing cyber forensics: what the social world has to offer. In 2015 AAAI Spring Symposium Series.

Cambridge Dictionary, 2019. Reconstruction, Available at: https://dictionary.cambridge.org/dictionary/english/reconstruction (Accessed: 28 July 2019)

Carrier, B. and Spafford, E.H., 2004, July. An event-based digital forensic investigation framework. In Digital forensic research workshop (pp. 11-13).

Casey, E., 2018. Clearly conveying digital forensic results. Digital Investigation, 24 1-3

Chung, H., Park, J. and Lee, S., 2017. Digital forensic approaches for Amazon Alexa ecosystem. Digital Investigation, 22, pp.S15-S25.

Cooper, P., Finley, G.T. and Kaskenpalo, P., 2010, June. Towards standards in digital forensics education. In Proceedings of the 2010 ITiCSE working group reports (pp. 87-95). ACM.

Daeid, Niamh Nic., Christian Cole, Michael Marra 2019 'Lords inquiry says forensic science is broken: here's how we can start to fix it' Available at: https://theconversation.com/lords-inquiry-says-forensic-science-is-broken-heres-how-we-can-start-to-fix-it-116456 (Accessed: 28 July 2019)

Gov.uk, 2019. 'Insufficient funding for forensic science puts justice at risk' Available at: https://www.gov.uk/government/news/insufficient-funding-for-forensic-science-puts-justice-at-risk

Grobler, M., 2010. Digital forensic standards: international progress.

Horsman, G., 2018. Framework for Reliable Experimental Design (FRED): A research framework to ensure the dependable interpretation of digital data for digital forensics. Computers & Security, 73, pp.294-306.

Horsman, G., 2019. Formalising investigative decision making in digital forensics: Proposing the Digital Evidence Reporting and Decision Support (DERDS) framework. Digital Investigation, 28, pp.146-151.

Horsman, G., 2019b. Raiders of the lost artefacts: championing the need for digital forensics research. Forensic Science International: Reports.

Horsman, G., 2019d. Tool testing and reliability issues in the field of digital forensics. Digital Investigation, 28, pp.163-175.

Journal Pre-proof

Journal Pre-proof

Kebande, V.R. and Ray, I., 2016, August. A generic digital forensic investigation framework for internet of things (iot). In 2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud) (pp. 356-362). IEEE.

Kokshoorn, B., Blankers, B.J., de Zoete, J. and Berger, C.E., 2017. Activity level DNA evidence evaluation: on propositions addressing the actor or the activity. Forensic science international, 278, pp.115-124.

Lillis, D., Becker, B., O'Sullivan, T. and Scanlon, M., 2016. Current challenges and future research areas for digital forensic investigation. arXiv preprint arXiv:1604.03850.

Martini, B. and Choo, K.K.R., 2012. An integrated conceptual digital forensic framework for cloud computing. Digital Investigation, 9(2), pp.71-80.

MD5., 2019 ‘Downloads’ Available at: https://vfc.uk.com/downloads/ (Accessed: 28 July 2019)

Mohay, G., 2005, November. Technical challenges and directions for digital forensics. In First International Workshop on Systematic Approaches to Digital Forensic Engineering (SADFE'05) (pp. 155-161). IEEE.

Nidhra, S. and Dondeti, J., 2012. Black box and white box testing techniques-a literature review. International Journal of Embedded Systems and Applications (IJESA), 2(2), pp.29-50.

Page, H., Horsman, G., Sarna, A. and Foster, J., 2019c. A review of quality procedures in the UK forensic sciences: What can the field of digital forensics learn?. Science & Justice, 59(1), pp.83-92.

Petroni Jr, N.L., Walters, A., Fraser, T. and Arbaugh, W.A., 2006. FATKit: A framework for the extraction and analysis of digital forensic data from volatile system memory. Digital Investigation, 3(4), pp.197-210.

Pollitt, M., Casey, E., Jaquet-Chiffelle, D. and Gladyshev, P., 2018. A framework for harmonizing forensic science practices and digital/multimedia evidence. Available at)(Accessed 16 May 2018) https://www. nist. gov/sites/default/files/documents/2018/01/10/osac_ts_0002. pdf Date.

Science and Technology Select Committee, Forensic science and the criminal justice system: a blueprint for change. 3rd Report of Session 2017-19 - published 1 May 2019 - HL Paper 333

Sommer, P., 2011. Certification, registration and assessment of digital forensic experts: The UK experience. digital investigation, 8(2), pp.98-105.

Sommer, P., 2018. Accrediting digital forensics: what are the choices?.

Journal Pre-proof

Sunde, N. and Dror, I.E., 2019. Cognitive and human factors in digital forensics: Problems, challenges, and the way forward. Digital Investigation, 29, pp.101-108.

The Guardian, 2019 ‘Police outsource digital forensic work to unaccredited labs’ Available at: https://www.theguardian.com/uk-news/2018/feb/12/police-outsource-digital-forensic-work-to-unaccredited-labs (Accessed: 28 July 2019)

Valjarevic, A. and Venter, H.S., 2012, August. Harmonised digital forensic investigation process model. In 2012 Information Security for South Africa (pp. 1-10). IEEE.

Journal Pre-proof

Journal Pre-proof

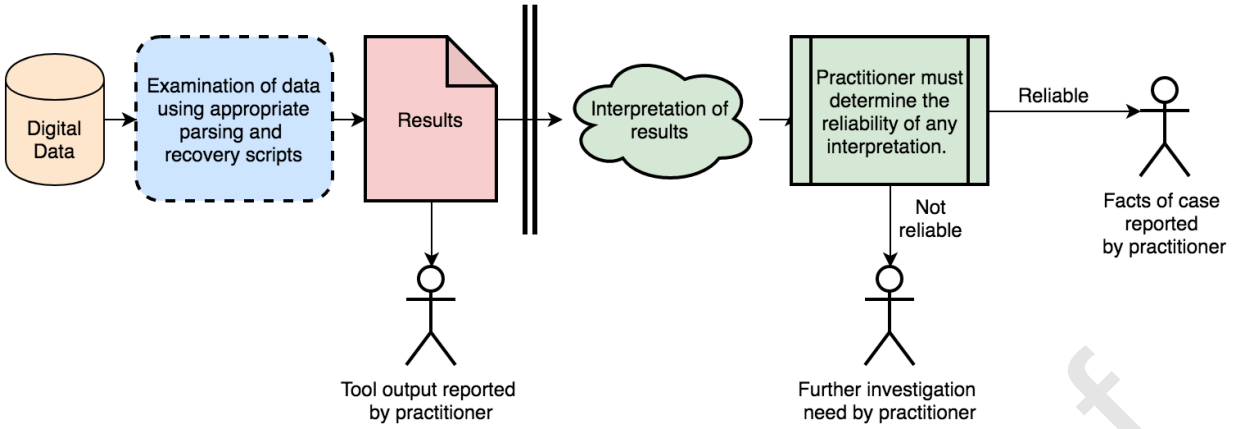

Figure 1: Reporting tool output Vs reporting the facts of a case.

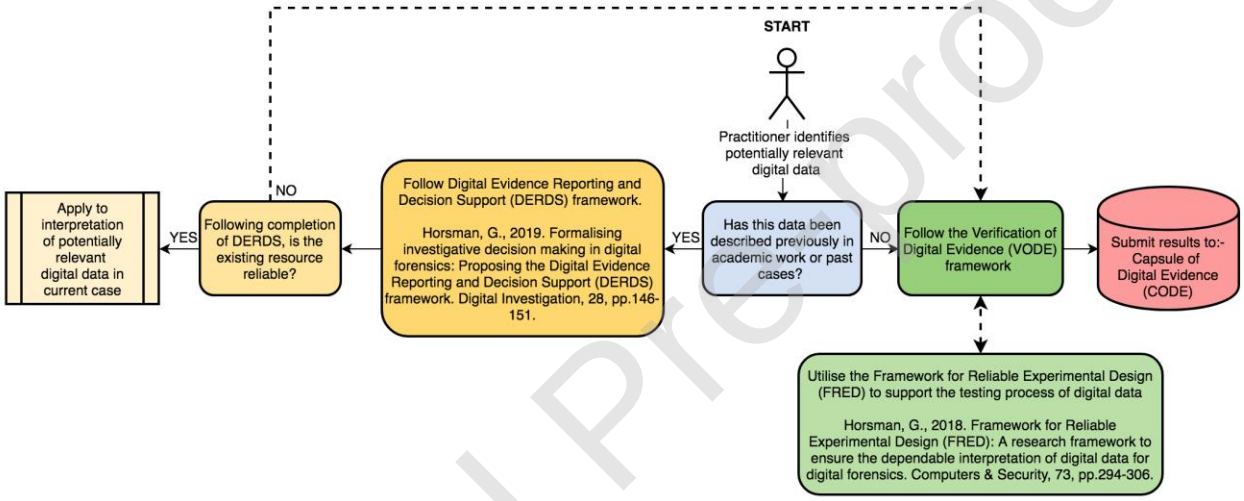

Figure 2: A road map of connecting frameworks (due to size, a higher quality image has been submitted as a separate file.)

Journal Pre-proof

Journal Pre-proof

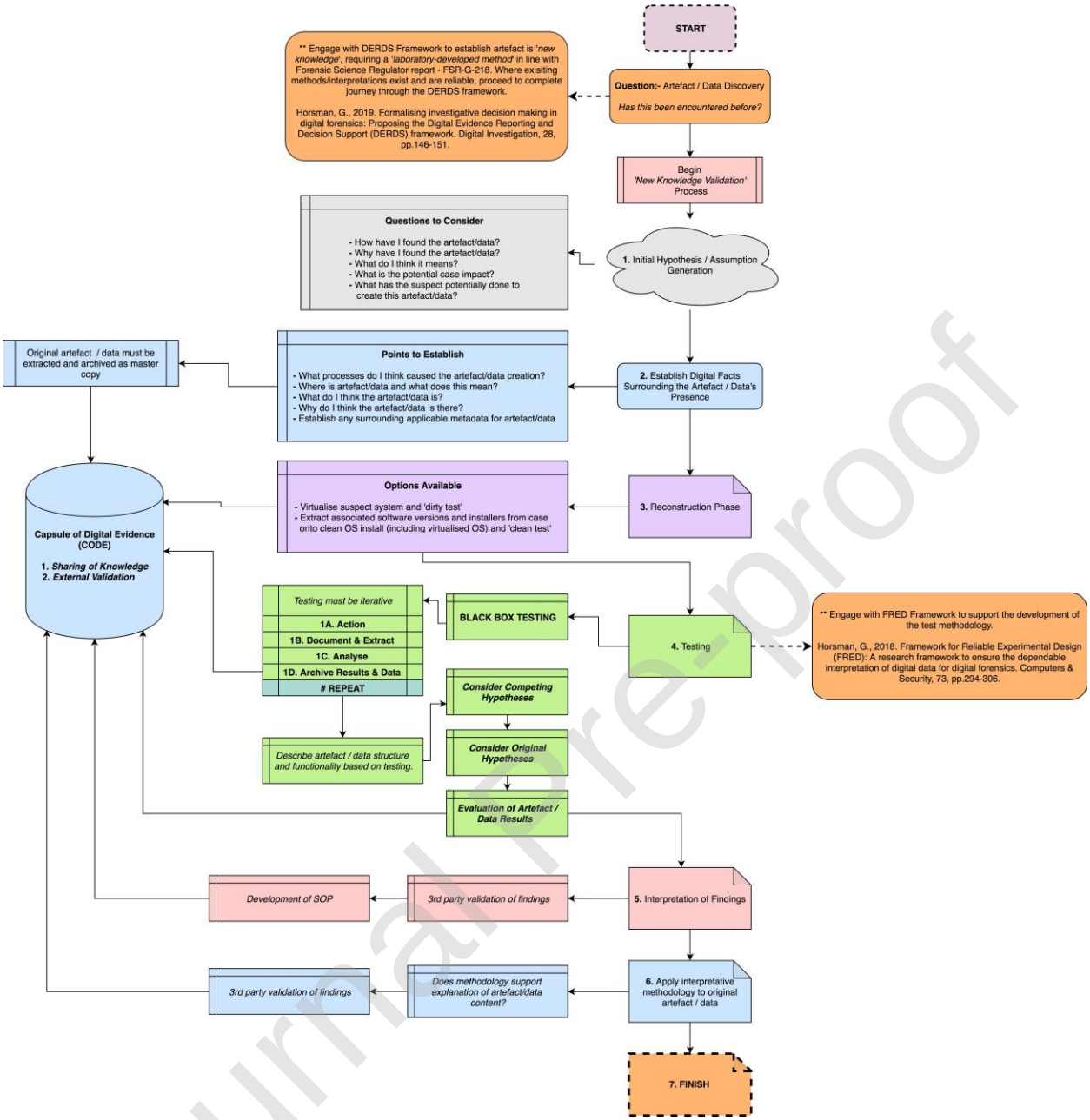

Figure 3: The Verification of Digital Evidence (VODE) framework (due to size, a complete image has been submitted as a separate file.)

Journal Pre-proof

Journal Pre-proof

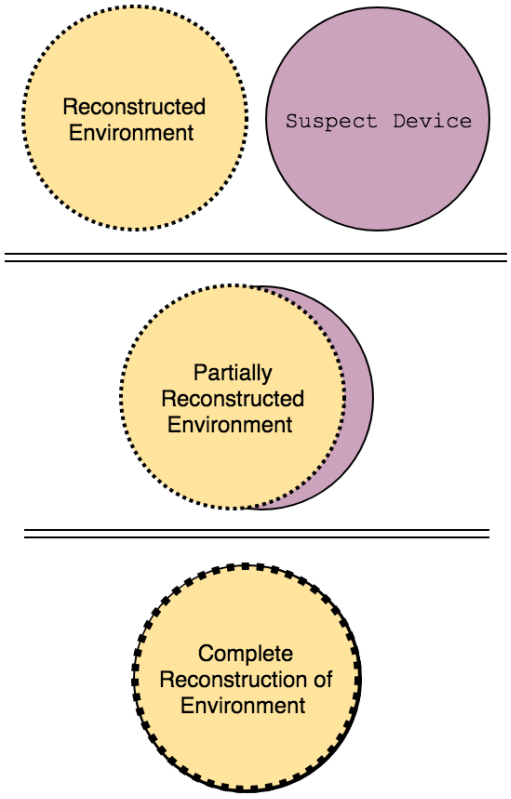

Figure 4: The stages of reconstruction

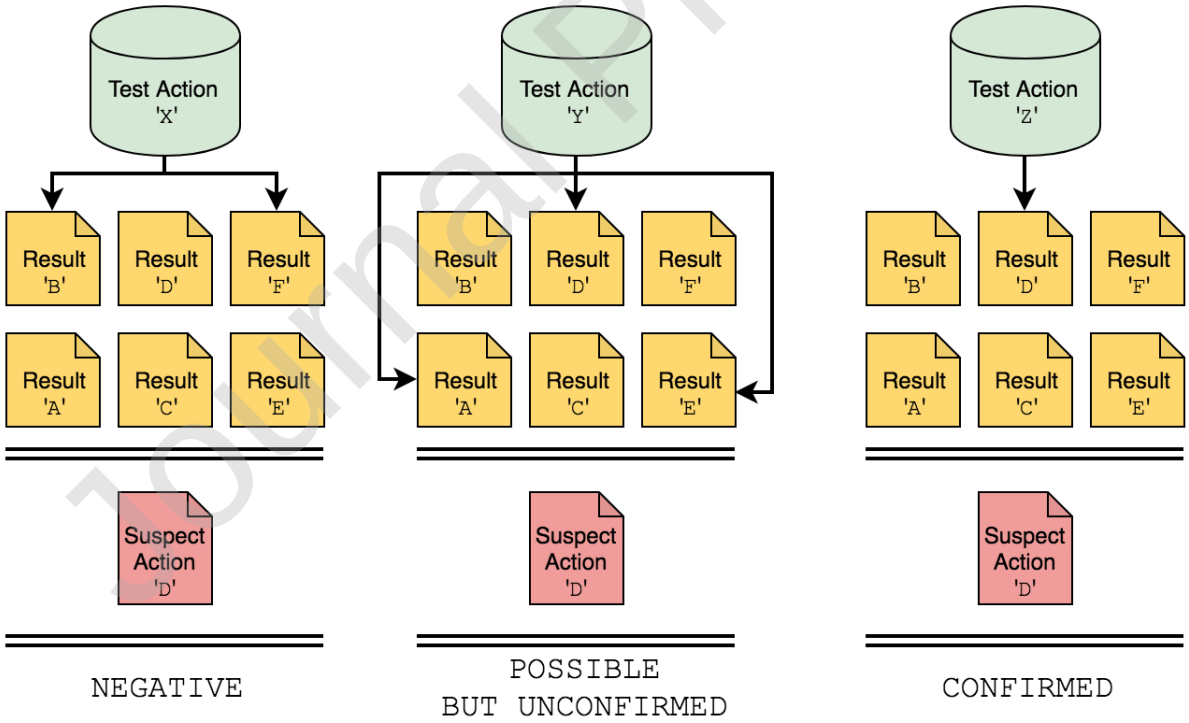

Figure 5: Testing outcomes and their meaning

Journal Pre-proof

Journal Pre-proof

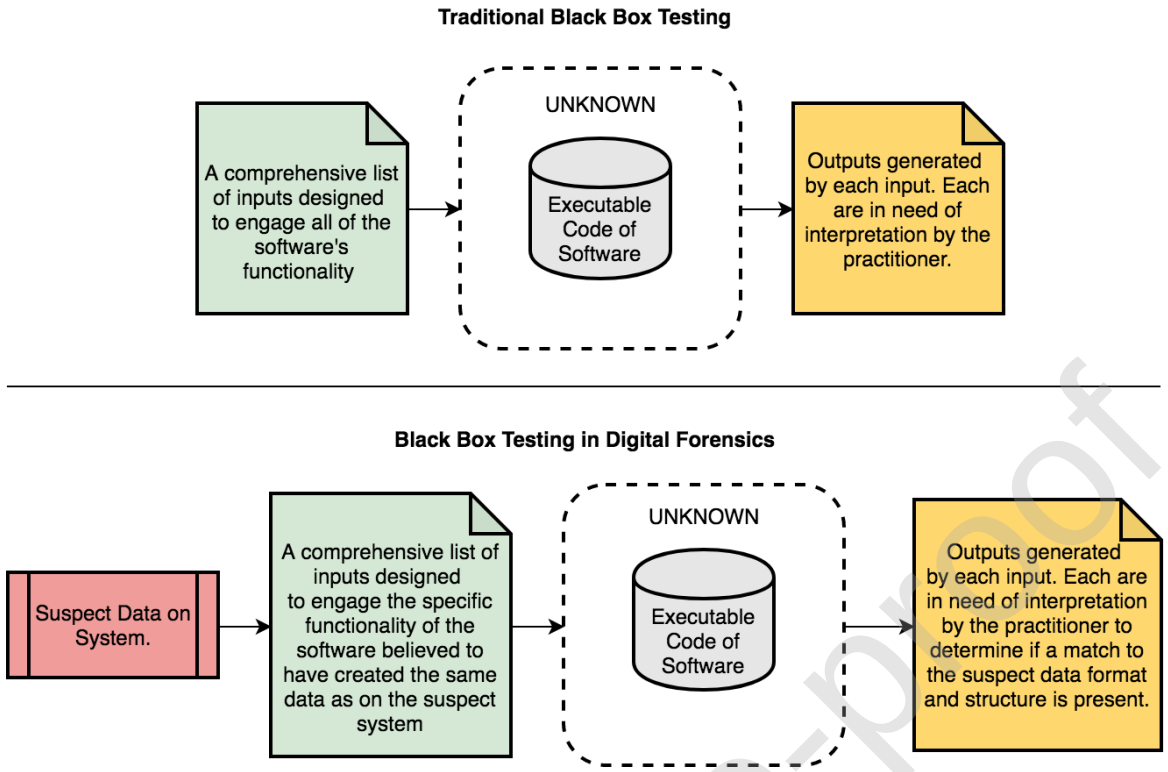

Figure 6: Black box testing in digital forensics

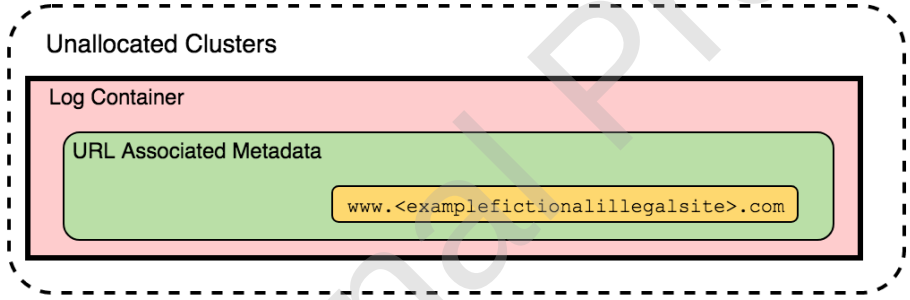

Figure 7: Example URL in unallocated.

Journal Pre-proof